

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 28, 2001, 19:50:29 : Search time 10.4 seconds

(without alignments)  
494.070 Million cell updated/sec

Title: US-09-516-052-2\_COPY\_28\_177

Sequence: 1 REUDGYMPLAVIRIMFKTL.....YGYGLDQSMWNGGRYUN 150

Scoring table:

BLASTSUM62

Gapop 10.0, Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_39.\*

Listing first 45 summaries

Result No. Score Query length DB ID Description

1 364 45 179 1 CHEA\_MAIZE P25209 GAI WAVE (M

2 324 5 40 4 CHEA\_PETMA P25210 FOLLOWING

3 322 5 40 1 CHEA\_CHICK F25207 GALLUS GALL

4 322 5 40 1 CHEA\_HUMAN F25208 HOMO SAPIEN

5 322 5 40 1 CHEA\_MOUSE F25209 MUS MUSCULA

6 289 34 8 HAV3\_YEAST F13434 SACCAROMYCE

7 279 5 34 8 HAV3\_SCHPO F13434 SACCAROMYCE

8 269 34 8 HAV3\_KCILA F13434 SACCAROMYCE

9 224 5 27 9 HAV3\_XENLA F13434 SACCAROMYCE

10 151 5 20 1 TRAC\_HUMAN F01658 HOMO SAPIEN

11 158 5 19 7 DR1\_AKATH F49592 ARABIDOPSIS

12 94 11 7 HAF2\_ARCFU F28779 ARCHAEOLLOB

13 94 11 7 HAF1\_ARCFU F28779 ARCHAEOLLOB

14 92 11 4 HANA\_THET1 F25659 LACTOCOCCUS

15 89 11 4 HAV1\_PYST F25659 LACTOCOCCUS

16 88 10 9 HMT1\_METH F25643 METHANOCOCC

17 88 10 9 HAV1\_PYST F25659 LACTOCOCCUS

18 87 10 9 HAV2\_METH F48783 METHANOCOCC

19 86 10 4 HAV2\_METH F48783 METHANOCOCC

20 84 10 4 HAV2\_METH F48783 METHANOCOCC

34 74 9 2 446 1 CRT\_HOMO G17279 HOMO SAPIEN  
35 74 9 2 2190 1 CCAD\_CHICK G17279 HOMO SAPIEN  
36 73 5 9 1 649 1 P147\_HUMAN G14444 HOMO SAPIEN  
37 73 5 9 1 666 1 HAV2\_PYST G14444 HOMO SAPIEN  
38 73 5 9 1 345 1 LEU3\_LACTA G02143 LACTOCOCCUS  
39 73 5 9 1 443 1 HX43\_MOUSE F02841 MUS MUSCULA  
40 73 5 9 1 526 1 KIC1\_BOVIN F06394 BOV TAUROS  
41 72 5 9 0 481 1 BIRD\_STRPU F06651 STRONGYLOCE  
42 72 5 9 0 648 1 WBN\_HUMAN G15353 HOMO SAPIEN  
43 72 5 9 0 1366 1 CA21\_CANPA G16392 CANIS FAMILI  
44 72 5 9 0 1786 1 YCF1\_AKATH F56785 ARABIDOPSIS  
45 71 8 8 962 1 ARVC\_HUMAN G00192 HOMO SAPIEN

#### ALIGNMENTS

RESULT 1  
ID CHEA\_MAIZE STANDARD: PRT: 179 AA.  
AC P25209  
DT 01-MAY-1992 (Ref. 22, Created)  
DT 01-MAY-1992 (Ref. 22, Last sequence update)  
DT 01-OCT-1996 (Ref. 34, Last annotation update)  
DE CCAT-BINDING TRANSCRIPTION FACTOR SUBUNIT A (CHE-A) (NF-Y PROTEIN  
DE CHAIN B) (NF-YB) (CAAT-BOX DNA BINDING PROTEIN SUBUNIT B).  
GN NFY2.  
OS Zea mays (Maize).  
OC Euphorbiaceae: Vitellipalmaceae: Embryophyta: Tracheophyta: Spermatophyta:  
OC Magnoliophyta: Liliopsida: Poales: Poaceae: PACO clade: Panicoideae:  
OC Andropogoneae: Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92195809; PubMed=1549471;  
RA NF-Y, Mannervik P., Boett van Huijsduijnen R., Andre L.,  
RA Benoist C., Mathis D.,  
RT "Evolutionary variation of the CCAT-binding transcription factor  
NF-Y".  
KL Nucleic Acids Res. 20:1087-1091(1992).  
CC FUNCTION: STIMULATES THE TRANSCRIPTION OF VARIOUS GENES BY  
CC RECOGNIZING AND BINDING TO A CCAT MOTIF IN PROMOTERS. FOR  
CC EXAMPLE IN TYPE I COLLAGEN, ALBUMIN AND BETA-ACTIN GENES.  
CC SUBUNIT: HETEROMERIC TRANSCRIPTION FACTOR COMPOSED OF TWO  
CC COMPONENTS, A AND B, THAT ARE BOTH NEEDED FOR DNA BINDING.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- DOMAIN. CAN BE DIVIDED INTO THREE DOMAINS. THE WEAKLY CONSERVED A  
CC DOMAIN. THE HIGHLY CONSERVED B DOMAIN SUBUNIT TO BE INVOLVED IN  
CC SUBUNIT INTERACTION AND DNA BINDING, AND THE GLO-RICH C DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE CHE-A SUBUNIT FAMILY.  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed, altered by and for commercial  
CC or other purposes. For more information, see http://www.ebi.ac.uk/seqdb/doc/

[illegible][illegible]

```

01 RESIDUE:1AAVARIMMNAITGTCIAIATACACVAGCVSEISTETSPASRCHQENKI 110
02 61 ITAEITLWAKELKLNPNVYIVININYPELETRDGLALGERPSLKTOGNIG 117
03 I [|||||] ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
04 DB 111 INGEHILFAMSLTGFDSVEFKTLCKLR-----AMGE-----KGIG 150
05 RESULT 4
06 CHFA_HUMAN STANDARD PRT 207 AA.
07 AC CHFA_HUMAN STANDARD PRT 207
08 DT 01-MAY-1992 (Ref. 22, Created)
09 DI 01-JUL-1993 (Ref. 26, Last sequence update)
10 DT 01-OCT-2000 (Ref. 40, Last annotation update)
11 DE NCBI_ELF_TRANSCRIPTION_FACTOR_T_SUBUNIT_BETA [NF-Y E-ELF1 CHAIN B;
12 (NF-YB)] COACT-BINDING TRANSCRIPTION FACTOR SUBUNIT A (CBF-A) (COAT-
13 BOX DNA BINDING PROTEIN SUBUNIT B).
14 NFB OR HAP3.
15 OS Homo sapiens (Human) .
16 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
17 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
18 OX NCBI_Taxid=9606;
19 RN 111
20 RP SEQUENCE FROM N.A.
21 RX MEDLINE=9225328P; PubMed=1577736;
22 RA LA X.Y.; van Halbeekoven R., Mantovani P., Benoist C.O.,
23 Marais D.;
24 RT "Intron exon organization of the NF-Y genes: tissue-specific splicing
25 modifies an activation domain."
26 J Biol Chem 267:4984-4990(1992);
27 LN 121
28 RP SEQUENCE FROM N.A.
29 RX Badley Clarke J., Ting J.P.Y.;
30 PI submitted (XXXX-1992) via ProEMBL; Genbank; Trnpt dr abasos
31 LN 131
32 RP SEQUENCE FROM N.A.
33 RX MEDLINE=9219500P; PubMed=154471;
34 RA LI X.Y., Mantovani P., Hoef van Halbeekoven R., Andre G.;
35 RA Benoist C.O., Marais D.;
36 RT "Evolutionary variation of the COAT-binding transcription factor
37 NF-Y."
38 RL Nucleic Acids Res 20:1997-1999(1992)
39 CC -1- FUNCTION: STIMULATES THE TRANSCRIPTION OF VARIOUS GENES BY
40 RECOGNIZING AND BINDING TO A COAT MOTIF IN PROMOTERS. FOR
41 EXAMPLE IN TYPE I COLLAGEN, ALBUMIN AND BETA-ACTIN GENES.
42 CC -1- SUBUNIT: HEPEMEREP TRANSCRIPTION FACTOR COMPOSED OF TWO
43 COMPONENTS A AND B, THAT ARE BOTH NECESS FOR DNA BINDING.
44 CC -1- SPECIFICITY LOCATION: NUCLEAR.
45 CC -1- DOMAIN- CAN BE DIVIDED INTO THREE DOMAINS. THE WEAKLY CONSERVED A
46 DOMAIN, THE HIGHLY CONSERVED B DOMAIN THOUGHT TO BE INVOLVED IN
47 SUBUNIT INTERACTION AND RNA BINDING, AND THE OLD FIRM C DOMAIN.
48 CC -1- SIMILARITY: BELONGS TO THE CBF-A SUBUNIT FAMILY.
49 CC
50 DT This SWISS-PROT entry is copyright. It is produced through a collaboration
51 between the Swiss Institute of Bioinformatics and the EMBL, situated at
52 the European Bioinformatics Institute. There are no restrictions on its
53 use by non-profit institutions as long as its content is in no way
54 modified and this statement is not removed or changed by any other
55 entities requires a license agreement (see http://www.ebi.ac.uk/embl/
56 or send an email to licenses@ebi.slb.ch)
57 CC
58 EMHL: L06145; AAA59410 1; -
59 EMHL: X59710; CAA42730 1; AT_1N1T
60 PIR: S22917; S22917
61 HSSP: P19267; IBPM.
62 DR TRANSFAC; T00154; -
63 DR MM: 189904; -
64 DR InterPro; IPRO00047; -
65 DR Pfam; PF00808; CBFD_NFYH_HME; 1
66 DR PRINTS; PR00615; COACTSUBUNITA
67 DR PROSITE; PS00685; CBPA_NFYB; 1
68 TR Transcription regulation; DNA-binding; Activator; Nuclear protein.
69 FT DOMAIN 1 52 A DOMAIN.

```

[illegible]







001 Mammaliophyta, eudicotyledons, core eudicots, rosidae, eurosid 11:  
 002 Brassicales, Brassicaceae, Arabidopsis.  
 003 NCBI\_TaxID 3702:  
 004 [1]  
 005 SEQUENCE FROM N.A.  
 006 STRAIN-CV, COLUMBIA:  
 007 MEDLINE-95116118; PubMed-7816619;  
 008 Kurotori T., Yamamoto M.:  
 009 "Cloning of cDNAs from Arabidopsis thaliana that encode putative  
 010 protein phosphatase 2s and a human DRP-like protein by transfection  
 011 of a fusion yeast mutant."  
 012 Nucleic Acids Res. 22:4394-4401(1994).  
 013 [1- SIMILARITY: TO HUMAN PROTEIN DRP].  
 014  
 015 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 016 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 017 the European Bioinformatics Institute. There are no restrictions on its  
 018 use by non-profit institutions as long as its content is in no way  
 019 modified and this statement is not removed. Usage by and for commercial  
 020 entities requires a license agreement. See http://www.isb.ac.uk/ncbi/ncbi.html  
 021 or send an email to license@isb.ac.uk.  
 022  
 023 EMBL: D38110; BAA07288.1; --  
 024 InterPro: IPR000947;  
 025 Pfam: PF00808; GRAFTM: GRAFTM1; 1.  
 026 Trnseq: Trnseq: Nucleic acid protein.  
 027 SEQUENCE 159 AA; 17642 MW; 858C8C969B8F54FE CRC64;  
 028  
 029 Query Match 19.7%; Score 158.5; DB 1; Length 159;  
 030 Best Local Similarity 25.6%; Pred. No. 13e-07;  
 031 Matches 30; Conservative 35; Mismatches 43; Indels 9; Gaps 1;  
 032  
 033 2 EASYMPYANVLPIMSKTETSHAPISQAFPTTFSVSVISVITGFANFPGPQPKRTI 61  
 034 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 035 11 KEDASLPKATMKTIKEMLPPIVAVAPDAQDLIECCVEFTLVSSSESDNCKEDKRTI 70  
 036 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 037 62 TAEGLIMMSKSLKPTNYPVPLVITFNY-----PELEDSALPPSPSPFQ 109  
 038 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 039 71 APEHVALAVLIGSGEITLEVAAVYQNHVFMCTQKSVKWNFACMTFFFAARQ 127  
 040 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 041  
 042 RESULT 12  
 043 HAP1\_ARCFU STANDARD; PRT; 67 AA  
 044 028779;  
 045 30-MAY-2000 (rel. 39, Created)  
 046 30-MAY-2000 (rel. 39, Last sequence update)  
 047 01-OCT-2000 (rel. 40, Last annotation update)  
 048 PROBABLE ARCHAEOAL HISTONE A1-2.  
 049 HPAV1-2 OR AF1433.  
 050 Archaeoglobus fulgidus.  
 051 Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
 052 Archaeoglobus.  
 053 NCBI\_TaxID-2234.  
 054 [1]  
 055 SEQUENCE FROM N.A.  
 056 STRAIN-VC-16 / DSM 4104 / AF03 49558;  
 057 MEDLINE-98049443; PubMed-9383475;  
 058 Klerk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 059 Richardson K.A., Eadsen R.T., Cain M., Hickey F.K., Peterson J.D.,  
 060 Fleischmann R.D., Quakekush J., Lee N.H., Sutton G.G., Gill S.,  
 061 Kitzman S.F., Dougherty R.A., McKenney K., Adams M.D., Loftus B.,  
 062 Peterson S., Ketch C.I., McNeil L.K., Rodger J.H., Glaser A., Zhou L.,  
 063 Cotton M.D., Springs T., Williams J.F., McDonald L., Utterback T.,  
 064 Sadov P.W., Oslen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 065 Mason T.M., Olson G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 066 Venter J.C.:  
 067 "The complete genome sequence of the hyperthermophilic, sulphate-  
 068 reducing archaeon Archaeoglobus fulgidus."  
 069

070 Nature 390:364-370(1997).  
 071 -1- FUNCTION: BINDS AND COMPACT DNA (95 TO 150 BASE PAIRS) TO FORM  
 072 NUCLEOSOME-LIKE STRUCTURES THAT CONTAIN POSITIVE DNA SUPERCOILS  
 073 INCREASES THE RESISTANCE OF DNA TO THERMAL DENATURATION (BY  
 074 SIMILARITY).  
 075 -1- SIMILARITY: BELONGS TO THE HMF FAMILY OF ARCHAEOAL HISTONES.  
 076  
 077 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 078 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 079 the European Bioinformatics Institute. There are no restrictions on its  
 080 use by non-profit institutions as long as its content is in no way  
 081 modified and this statement is not removed. Usage by and for commercial  
 082 entities requires a license agreement. See http://www.isb.ac.uk/ncbi/ncbi.html  
 083 or send an email to license@isb.ac.uk.  
 084  
 085 EMBL: A6000999; AAB89751.1; --  
 086 HSP; P48781; IPIV.  
 087 IIGR: AF1493;  
 088 InterPro: IPR000947;  
 089 Pfam: PF00808; GRAFTM: GRAFTM1; 1.  
 090 Trnseq: Trnseq: Nucleic acid protein.  
 091 SEQUENCE 67 AA; 7252 MW; 52FA01086968D69 CRC64;  
 092  
 093 Query Match 11.7%; Score 94; DB 1; Length 67;  
 094 Best Local Similarity 35.9%; Pred. No. 0.029;  
 095 Matches 21; Conservative 21; Mismatches 26; Indels 2; Gaps 1;  
 096  
 097 7 MTIANVLEMEPTETTHAKSPDQVPTIGPVSFVTSVSTGTFANFPGPQPKRTI 66  
 098 111 111 111 111 111 111 111 111 111 111 111 111 111 111  
 099 4 LPMVAVLELFEA-CAPVSAVAVFVNFVLEIVDTAKKAVPLAKHSCKRTVADDI 61  
 100 111 111 111 111 111 111 111 111 111 111 111 111 111  
 101 67 LMAVS 71  
 102 111  
 103 62 KTAIS 66  
 104 111 111 111 111 111 111 111 111 111 111 111 111 111  
 105  
 106 RESULT 13  
 107 HAP1\_ARCFU STANDARD; PRT; 72 AA  
 108 029910;  
 109 30-MAY-2000 (rel. 39, Created)  
 110 30-MAY-2000 (rel. 39, Last sequence update)  
 111 01-OCT-2000 (rel. 40, Last annotation update)  
 112 PROBABLE ARCHAEOAL HISTONE A1-1.  
 113 HPAV1-1 OR AF0337.  
 114 Archaeoglobus fulgidus.  
 115 Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
 116 Archaeoglobus.  
 117 NCBI\_TaxID-2234.  
 118 [1]  
 119 SEQUENCE FROM N.A.  
 120 STRAIN-VC-16 / DSM 49558;  
 121 MEDLINE-98049443; PubMed-9383475;  
 122 Klerk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 123 Richardson K.A., Eadsen R.T., Cain M., Hickey F.K., Peterson J.D.,  
 124 Fleischmann R.D., Quakekush J., Lee N.H., Sutton G.G., Gill S.,  
 125 Kitzman S.F., Dougherty R.A., McKenney K., Adams M.D., Loftus B.,  
 126 Peterson S., Ketch C.I., McNeil L.K., Rodger J.H., Glaser A., Zhou L.,  
 127 Cotton M.D., Springs T., Williams J.F., McDonald L., Utterback T.,  
 128 Sadov P.W., Oslen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 129 Mason T.M., Olson G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 130 Venter J.C.:  
 131 "The complete genome sequence of the hyperthermophilic, sulphate-  
 132 reducing archaeon Archaeoglobus fulgidus."  
 133 Nature 390:364-370(1997).  
 134 -1- FUNCTION: BINDS AND COMPACT DNA (95 TO 150 BASE PAIRS) TO FORM  
 135 NUCLEOSOME-LIKE STRUCTURES THAT CONTAIN POSITIVE DNA SUPERCOILS  
 136 INCREASES THE RESISTANCE OF DNA TO THERMAL DENATURATION (BY  
 137 SIMILARITY).  
 138 -1- SIMILARITY: BELONGS TO THE HMF FAMILY OF ARCHAEOAL HISTONES.  
 139





Thu Nov 29 08:24:21 2001

us-09-516-052-2\_copy\_28\_177.rsp

Page 9

UY 67 LWAM 70  
1:  
Dbb 62 KLA1 65

Search completed: November 28, 2001, 19:53:55  
Job time: 206 sec

